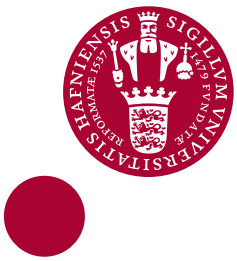

Development and clinical validation of a circulating tumor DNA test for the identification of clinically actionable mutations in nonsmall cell lung cancer

Liu, Liping; Liu, Han; Shao, Di; Liu, Zu; Wang, Jingjing; Deng, Qiuhua; Tang, Hailing; Yang, Haihong; Zhang, Yalei; Qiu, Yuan; Cui, Fei; Tan, Meihua; Zhang, Pan; Li, Zhilong; Liu, Jilong; Liang, Wenhua; Wang, Yuying; Peng, Zhiyu; Wang, Jian; Yang, Huanming; Mao, Mao; Kristiansen, Karsten; Ye, Mingzhi; He, Jianxing

Published in:
Genes, Chromosomes & Cancer

DOI:
10.1002/gcc.22522

Publication date:
2018

Document version
Publisher's PDF, also known as Version of record

Received: 22 September 2017 | Revised: 20 December 2017 | Accepted: 21 December 2017

DOI: 10.1002/gcc.22522

RESEARCH ARTICLE

WILEY

Development and clinical validation of a circulating tumor DNA test for the identification of clinically actionable mutations in nonsmall cell lung cancer

Liping Liu[1,2*] | Han Liu[3*] | Di Shao[3,4,5] | Zu Liu[3] | Jingjing Wang[3] | Qiuhua Deng[1,2] | Hailing Tang[1,2] | Haihong Yang[6] | Yalei Zhang[6] | Yuan Qiu[6] | Fei Cui[6] | Meihua Tan[3,4] | Pan Zhang[3] | Zhilong Li[3] | Jilong Liu[3] | Wenhua Liang[1,6] | Yuying Wang[4] | Zhiyu Peng[4] | Jian Wang[7,8] | Huanming Yang[7,8] | Mao Mao[4] | Karsten Kristiansen[5,7] | Mingzhi Ye[3,4,5] | Jianxing He[1,6]

[1]State Key Laboratory of Respiratory Diseases, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Disease, Guangzhou 510120, China

[2]The Translational Medicine Laboratory, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China

[3]BGI-Guangzhou Medical Laboratory, BGI-Shenzhen, Guangzhou 510006, China

[4]BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China

[5]Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, Copenhagen, DK 2200, Denmark

[6]Department of Thoracic Surgery, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China

[7]BGI-Shenzhen, Shenzhen 518083, China

[8]James D. Watson Institute of Genome Sciences, Hangzhou 310058, China

Correspondence
Jianxing He, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China.
Email: hejx@vip.163.com
or
Mingzhi Ye, BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen 518083, China.
Email: yemingzhi@genomics.cn

Funding information
National Key R&D Program of China, Grant Number: 2016YFC0905400; the Science and Technology Planning Project of Guangdong Province, People's Republic of China, Grant Numbers: 2007B031515017 and 2008A030201024; the Science and Technology Planning Project of Guangzhou, People's Republic of China, Grant Numbers: 2007Z1-E0111 and 2007Z3-E0261; the Guangzhou Health and Medical Collaborative Innovative Major Special Projects, Grant Number: 201400000001-2; special

Abstract

Molecular analysis of potentially actionable mutations has become routine practice in oncological pathology. However, testing a wide range of oncogenes and mutations can be technically challenging because of limitations associated with tumor biopsy. Circulating tumor DNA (ctDNA) is a potential tool for the noninvasive profiling of tumors. In this study, we developed a next-generation sequencing (NGS)-based test for the detection of clinically relevant mutations in ctDNA and evaluated the feasibility of using this ctDNA NGS-based assay as an alternative to tissue genotyping. Tissue and matched blood samples were obtained from 72 patients with advanced nonsmall cell lung cancer (NSCLC). NGS-based testing was performed using plasma cell-free DNA (cfDNA) samples of all 72

*These authors contribute equally to this study.

foundation for High-level Talents of Guangdong, Grant Number: 2016TX03R171; Guangzhou Science and Technology Project, Grant Number: 201400000004-5; Pearl River Nova Program of Guangzhou, Grant Number: No. 201506010065; and the Natural Science Foundation for Distinguished Young Scholars of Guangdong Province, Grant Number: No. 2014A030306013

patients as well as tumor DNA samples of 46 patients. Of the remaining 26 patients, tDNA was tested by amplification refractory mutation system PCR (ARMS-PCR) because of insufficient tissue sample or quality for NGS. Of the 46 patients who had tDNA and cfDNA NGS performed, we found 20 patients were concordant between tDNA and ctDNA alterations and 21 sample pairs were discordant because of additional alterations found in tDNA. Considering all clinically relevant alterations, the concordance rate between tDNA and ctDNA alterations was 54.9% with a sensitivity of 53.2% and a specificity of 75.0%. Our findings demonstrate that targeted NGS using cfDNA is a feasible approach for rapid and accurate identification of actionable mutations in patients with advanced NSCLC, and may provide a safe and robust alternative approach to tissue biopsy.

KEYWORDS
actionable mutation, ctDNA, NGS, NSCLC

1 | INTRODUCTION

Lung cancer is the leading cause of cancer-related mortality in the world.[1] In China, the mortality of lung cancer increased dramatically in the last three decades.[2] Targeted therapeutics have shown considerably better clinical efficacy compared with standard therapy in patients with nonsmall cell lung cancer (NSCLC) and an oncogenic mutation.[3–6] Of these cases, most were adenocarcinoma in which an increasing number of oncogenic driver mutations have been discovered in recent years. In the latest version of the National Comprehensive Cancer Network (NCCN), *EGFR*, *ALK*, and *ROS1* testing were recommended before making therapy decisions. Additionally, emerging targeted drugs were listed for genetic alterations including *BRAF* V600E, *MET* amplification or the exon 14 skipping mutation, *RET* rearrangements, and *HER2* mutations.[7] As there is a steady increase in the number of targeted genes through existing genotype-selected trials, the capacity for multiplexing detection will become crucial and even indispensable in the near future. Thus, a platform such as next-generation sequencing (NGS) that enables the detection of multiple tumor-specific mutations in a single assay is urgently needed.

Tissue biopsy is the gold standard for molecular testing, yet poor quality and inadequate quantity of tissue obtained makes the test inaccessible to approximately 20–30% patients when tissue is obtained from core needle biopsies.[8] In contrast, liquid biopsy could address this issue given the relative ease of blood collection that is much safer and less stressful for the patient. Moreover, there are several other advantages of testing circulating tumor DNA (ctDNA) over tissue biopsy: (1) rebiopsy is not always feasible with tissue, but ctDNA is accessible during the entire disease course; (2) identify treatment resistance by the detection of resistant clones that could be used to guide the subsequent treatment at disease progression and explore mechanism of novel resistance; and (3) regarding cost effectiveness and turnaround time, liquid biopsy is faster and can cost significantly less than tissue biopsy considering the full cost in obtaining tissue samples and the potential extra costs caused by biopsy complications.[9] With growing evidence demonstrating the reliability of ctDNA in genetic profiling, the Food and Drug Administration (FDA) from the United States approved the Cobas EGFR Mutation Test v2 as a companion diagnostic kit for NSCLC in June 2016.

A meta-analysis concluded that ctDNA testing showed high diagnostic accuracy; however, the sensitivity varies considerably depending on the platform and detection method.[10] Further, challenges remain in using ctDNA for liquid biopsy testing because of the inherently ultra-low concentration of ctDNA found in peripheral blood.[11] Thus, NGS sequencing procedures and variant calling strategies must be improved to detect low abundancy mutations while avoiding false positive results. Previous studies have specifically focused on the detection of a single biomarker or a few hotspot mutations such as *EGFR* mutations using single nucleotide variants (SNVs) or insertion/deletions (indel). The purpose of this study was to evaluate the detection performance of four types of variants in ctDNA on an ion proton platform using matched tumor DNA (tDNA) and cell-free DNA (cfDNA) from plasma of patients with advanced NSCLC.

2 | MATERIALS AND METHODS

2.1 | Ethics statement and patients

The study was approved by the Institutional Review Board of the Beijing Genomics Institute. All 72 patients who participated in this study provided written informed consent for the use of blood and lung tumor tissue. Patients were eligible for the study if they were diagnosed with advanced NSCLC using tissue biopsy and extra blood and tumor tissue samples were available for NGS testing (Figure 1). Patients who received treatment prior to resection or biopsy were excluded. Most participating patients were highly suspected to have genetic alternations based on immunohistochemistry (IHC) results. All blood samples were obtained during the week before surgery or biopsy collection and were immediately processed to isolate plasma. Tissue samples were collected from the primary site during surgery or biopsy. Relevant clinical and pathological information of the subject, including smoking history and metastatic sites, were obtained from the electronic health record. Cell line genomic DNA was purchased from the American Type Culture Collection (ATCC, Manassas, VA). Confirmation of cell line genomic DNA was performed by allele-specific PCR and Sanger sequencing.

2.2 | Tumor tissue genotyping

DNA was extracted from tumor tissue using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany). Amplification refractory mutation system PCR (ARMS-PCR) was used to detect *EGFR* and *KRAS* using

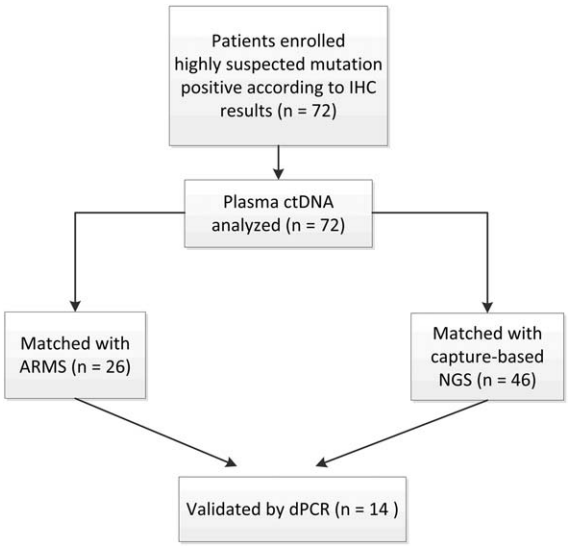

FIGURE 1 Overview of the study design for the detection of mutations in tumor tissue and plasma samples of patients with advanced lung cancer

commercial kits from AmoyDx (Xiamen, China), and fluorescence *in situ* hybridization (FISH) was used to detect ALK mutations using probes from Abbott (Shields, IL). For 63.9% of patients (46/72) who had sufficient DNA samples, tDNA NGS was performed using a previously validated targeted NGS assay.[12] Hybridization-based capture from 13 introns and 436 exons in 145 cancer-related genes including actionable recurrent rearrangements and amplifications was performed. Hybrid capture libraries were then sequenced to a read depth [(>)×200 coverage] using Ion Proton Sequencers (Thermo Fisher, Waltham, MA). A clinically validated bioinformatics pipeline named 'Otype' was used to detect clinically relevant genomic alterations.[12]

2.3 | Plasma sample genotyping

cfDNA was extracted from 4 to 5 mL of plasma using the QIAamp Circulating Nucleic Acid Kit (Qiagen) according to the manufacturer's instructions. The input cfDNA for the assay is 30 ng to achieve the expected assay performance. Sequencing libraries were prepared and hybridized to the same panel as that used for the tumor sample. The libraries were then quantified and qualified using a 2100 Bioanalyzer with High Sensitivity DNA 1000 Analysis Kit (Agilent, Santa Clara, CA). Finally, libraries were amplified by emulsion PCR on the Ion OneTouch 2 Instrument and loaded onto the Ion Proton Sequencer.

We developed a new Variation Hotspot Validation Toolkit (VHVT) for the detection of four types of actionable variants in plasma, specifically SNVs, indels, rearrangements, and copy number variations (CNVs). First, reference sequences aimed at the hotspot mutations were assembled, then reads were mapped to the new assembled reference to precisely distinguish the supporting reads. Moreover, log odds (LOD) and Poisson mathematical model were integrated to control sequencing errors. As a result, VHVT can achieve a limitation of detection at 0.01% with sensitivity and specificity above 95% and 99%, respectively. Detailed information regarding this software can be found in an ASCO Abstract we reported previously.[13] Rearragements were detected using tiled probes that allow discovery of known rearrangements. Hybrid baits covered all potential rearranged regions and were designed based on hotspot breakpoints that reported by COSMIC. All introns frequently rearranged were contained in the panel with high density. The assay contains 57 773 different baits in total. Of these, 2522, 472, and 2237 baits were designed to capture RET, ALK and ROS1 fusions, respectively. For ALK rearrangement, the baits focus on intron 19, exon 19 and exon 20. For the detection of CNV, we used a *Z*-score and hierarchical clustering method that have been reported previously.[14,15]

2.4 | Mutation validation by digital PCR

Digital PCR (dPCR) was carried out to confirm the identification of hotspot mutations in tDNA that were absent in plasma cfDNA via NGS. Mutant allele frequency was measured using the QuantStudio 3D Digital PCR System (Thermo Fisher) and QuantStudio 3D Digital PCR 20K Chip Kit v2 in accordance with the manufacturer's instructions. Briefly, 5–10 ng of cfDNA was added to dPCR reaction mixtures containing primers and probe. The dPCR reaction was performed with the following conditions: 1 cycle of 96°C for 10 min followed by 39 cycles of 56°C for 2 min and 98°C for 30 s, 1 cycle of 60°C for 2 min, and 1 cycle of 10°C for 10 min. dPCR data were analyzed using QuantStudio 3D Analysis Suite Software (Thermo Fisher).

2.5 | Statistical analysis

A normal distribution assumption for continuous variables was tested by the Shapiro-Wilk normality test. Continuous variables were expressed as the mean and SD, or if not normally distributed as the interquartile range (IQR). Correlations between continuous variables were assessed by the Spearman correlation coefficient. The relationship between patient clinical characteristics and concordance status was measured by chi-square test. Differences in two independent samples of continuous variables were evaluated by the Wilcoxon Rank Sum test or the Kolmogorov-Smirnov test, and when comparing more than two independent samples, the Kruskal-Wallis test was used.

Using gene mutations identified in tumor tissues as reference, we tabulated true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP). Matched tissue and plasma samples carrying the same mutations were classified as TP, while matched tissue and plasma samples without somatic mutations were TN. Gene mutations found in tissue but not in plasma samples were defined as FN, and gene mutations found in plasma samples but not in tissue were FP. Measures of TP, TN, FN, and FP were used to calculate sensitivity, specificity, and concordance rate.

All statistical analyses were two sided and a $P < .05$ was considered statistically significant. Analyses were performed using R 3.3.

3 | RESULTS

3.1 | Patient characteristics

Tissue and matched blood samples were obtained from 72 patients with advanced NSCLC. Demographic and clinical characteristics of the 72 patients are listed in Table 1. Forty-four (61.1%) patients were male, and 39 (54.2%) were nonsmokers. Most patients had an

TABLE 1 Patient characteristics

Characteristic	All patients $n = 72$	Patients with paired samples (%)	$n = 46$	(%)
Gender				
Male	44	61.1	30	65.2
Female	28	38.9	16	34.8
Age (years)				
Median	59		57.5	
Range	40–83		39–83	
Smoker				
No	39	54.2	24	52.2
Ever	29	40.3	18	39.1
NA	4	5.6	4	8.7
Stage				
IIIa	18	25.0	6	13.0
IIIb	5	6.9	4	10.9
IV	35	48.6	22	47.8
NA	14	19.4	14	30.4
Histology				
ADC	63	87.5	40	87.0
SCC	3	4.2	2	4.3
NOS	6	8.3	4	8.7

Abbreviations: ADC, adenocarcinoma; NA, not available; NOS, not otherwise specified; SCC, squamous cell carcinoma.

adenocarcinoma histological subtype ($n = 63$, 87.5%), and the remaining patients ($n = 9$) had a diagnosis of either a squamous cell histological subtype ($n = 3$, 4.2%) or NSCLC-not otherwise specified ($n = 6$, 8.3%) (Table 1). At the time of diagnosis, 58 patients whose cancer was staged were at either stage III (31.9%) or stage IV (48.6%).

3.2 | Assay performance assessed by cell lines

Cell lines with known variants H1975 (*EGFR* L858R and T790M), PC-9 (*EGFR* p.E746-A750del), H2228 (*ALK-EML4*), HCC78 (*ROS1_SLC34A2*), and SNU (*MET* amplification) were diluted to assess the analytical sensitivity and limit of detection (LOD) for the different variant types (SNV, indel, CNV, and rearrangement). We mixed DNA from mutated cell lines with DNA from wildtype cell lines to derive various target allele frequencies (30%, 10%, 5%, 1%, and 0.5%). Based on the established sequencing coverage (×800-×1000), the LOD was determined using a mathematical model and then validated on DNA dilutions derived from cell lines (Supporting Information Table S1). The LOD was defined as a mutation frequency that can be successfully detected greater than 95% chance. For SNVs and indels, we determined the analytical sensitivity as 100% at a LOD of 1%. At an allele frequency of 0.5%, 15/16 (93.8%) SNVs and 6/6 (100%) indels were detected. Regarding rearrangements and CNVs, we detected all mutations with an allele frequency ≥10% (Supporting Information Table S1). The measured allele frequency was in accordance with the expected allele frequency. We found that the assay showed a good inter-run reproducibility when assessed using eight samples harboring three mutations with an allele frequency of 1% (Supporting Information Table S2). While reads containing SNVs and short indels can be mapped to the right position using Gatk-like programs, a program designed for germline mutation detection and widely used for somatic mutation detection, reads containing long indels, such as EGFR EX19 deletions are often wrongly mapped especially when there are deletions near the ends of the reads. Thus, Gatk did not recognize these reads, consequently underestimating the mutation allelic frequency, and even missed out indels when supporting reads were rare. These findings demonstrated that a high analytical performance was achieved using VHVT,[13] which unlike other toolkits, shows an outstanding capacity to discriminate reads, and thus, enables the detection of actionable *EGFR* EX19 deletions.

3.3 | cfDNA concentration and clinical relevance

Of the 72 plasma samples collected, the median cfDNA concentration in plasma was 15.32 ng/mL (IQR = 9.26 ng/mL). cfDNA concentration was not significantly associated with metastasis status, although we observed that cfDNA concentrations were slightly higher in patients with metastasis (median concentration of 13.25 vs. 16.98 ng/mL, respectively) (Supporting Information Table S3). We also found that cfDNA concentrations were not significantly associated with gender. The median concentration was 13.16 ng/mL in males and 16.66 ng/mL in females ($P = .063$).

3.4 | Somatic mutations detected in tumor samples by NGS

Of the 72 plasma samples obtained in our study, 46 matched tumor samples were available for mutation detection with NGS. Other samples were either unobtainable or of an insufficient quantity or quality after performing ARMS testing in the hospital. Of these 46 tumor DNA samples that were successfully sequenced on the Ion Proton platform, we achieved an average sequencing depth of 372× (Supporting Information Figure S1). Mutations were identified in 89% of cases (41/46) and included SNVs, indels, CNVs, and rearrangements. We identified 16 variants in 9 genes, including *EGFR* mutations in 17 samples, *KRAS* mutations in 9 samples, *ALK* fusions in 5 samples, *MET* copy number gains in 3 samples, and *ERBB2*, *NRAS*, *RET*, *BRAF*, and *PIK3CA* mutations in 1 sample each. Six samples had two variants, and one sample had three variants. The mutation profiles of patients with matched tumor and plasma NGS results are shown in Tables 2 and 3.

3.5 | Somatic mutations detected in plasma samples by NGS

Seventy-two plasma DNA samples were successfully sequenced on the Ion Proton platform. The average sequencing depth was 938× (Supporting Information Figure S2). Thirty-nine of the 72 patients had ≥1 potentially actionable mutation. No somatic variants were detected in the plasma cfDNA of 33 patients. Variants were detected in 9 genes the most prevalent of which was *EGFR* that was found in 28 samples and accounting for 71.8% of samples with detected variants. Among actionable mutations, we found that *EGFR* L858R mutations occurred slightly more frequently than *EGFR* exon 19 deletions being found in

TABLE 2 Concordant mutations detected by NGS in matched tissue and plasma samples

Sample ID	Tumor stage	cfDNA concentration (ng/mL)	Mutation detected by NGS in tissue (MAF)	Mutation detected by NGS in plasma (MAF)	Sequencing coverage in plasma
1	IV	5.5	*KRAS* G12C (13.9%)	*KRAS* G12C (1.0%)	492.99
2	IV	66.1	*BRAF* V600E (32.1%)	*BRAF* V600E (14.2%)	138.54
3	IV	11.6	*ALK-EML4*	*ALK-EML4*	1223.82
4	IV	10.77	*ALK-EML4*	*ALK-EML4*	1118.1
5	IV	41.58	*EGFR* L858R (13%)	*EGFR* L858R (2.2%)	962.39
6	IV	9.02	*EGFR* L858R (2%)	*EGFR* L858R (10.3%)	750.1
7	IV	13.42	*EGFR* exon 19 del (17.2%)	*EGFR* exon 19 del (4.2%)	250.9
8	NA	21.36	*EGFR* exon 19 del (50.2%)	*EGFR* exon 19 del (7.1%)	1120.9
9	IV	13.08	*EGFR* L858R (59.6%)	*EGFR* L858R (1.2%)	1161.76
10	NA	18.73	*EGFR* exon 19 del (81%)	*EGFR* exon 19 del (24.8%)	664.91
11	IV	15.87	*RET-KIF5B*	*RET-KIF5B*	1154.5
12	IIIb	12.22	*ALK-EML4*	*ALK-EML4*	883.74
13	IV	8.88	*EGFR* L861R (68.6%)	*EGFR* L861R (9.4%)	1098.38
14	IIIa	17.6	*KRAS* G12C (59.1%)	*KRAS* G12C (18.8%)	1168.63
15	IIIb	8.72	*KRAS* G12C (48.3%)	*KRAS* G12C (12.6%)	1007.27
16	IIIb	18.6	*EGFR* exon 20 ins (30.8%)	*EGFR* exon 20 ins (3.3%)	1007.27
17	NA	7.38	*EGFR* exon 19 del (54.5%)	*EGFR* exon 19 del (19.5%)	490.58
18	IV	25.65	*EGFR* L858R (17.7%)	*EGFR* L858R (0.4%)	1234.05
19	NA	20.71	*EGFR* exon 19 del (51%)	*EGFR* exon 19 del (5.2%)	1329.51
20	IV	18.8	*EGFR* exon 19 del (22.3%)	*EGFR* exon 19 del (0.6%)	1106.13
42	IIIa	13.2	NEG	NEG	1374.76
43	IIIa	4.3	NEG	NEG	543.32
44	IIIa	21.2	NEG	NEG	786.93
45	IIIa	8.7	NEG	NEG	988.87
46	IV	12.77	NEG	NEG	1087.14

Abbreviations: MAF, minor allele frequency; NA, not available; NEG, negative.

16 and 11 samples, respectively. We found that three patients had the *EML4-ALK* rearrangement, which is sensitive to the ALK tyrosine-kinase inhibitor (TKI) crizotinib, and one patient had a *RET* rearrangement, which is sensitive to cabozantinib. In addition, we detected *KRAS* mutations in six plasma samples and *BRAF* V600E and *EGFR* L861R in one sample each. The *KRAS* mutations included G12A, G12C, and G12R, variants associated with original resistance to EGFR TKI in lung cancer. We found no co-occurring mutations in any of the plasma samples tested.

3.6 | Concordance between tumor and plasma results

We compared clinically relevant genomic alterations for 46 patients (68%) in our 72-patient cohort who had both tDNA and cfDNA NGS performed (Figure 2). We found that 20 patients have concordant tDNA and cfDNA alterations, and 5 patients have no alterations in either tDNA or plasma cfDNA. Of these 20 concordant sample pairs, the average allele frequency of mutations in tDNA was significantly greater than that found in plasma cfDNA (26.9% vs. 7.4%, $P < .001$). The remaining 21 sample pairs were discordant because additional alterations were found in tDNA. We found a significant positive correlation between tumor and cfDNA variant allele frequency (Spearman $r = 0.44$, $P = .027$).

Considering all clinically relevant alterations, the concordance rate between tDNA and cfDNA alterations was 54.9% with a sensitivity of 53.2%, specificity of 75.0%, and a plasma positive predictive value of 96.2%. For tDNA and plasma cfDNA from the 19 patients with metastasis, the alteration concordance rate was 65%. On the other hand, the 19 samples from primary patients had a mutation concordance of 50%. We also studied concordance for the most frequent alterations for all 72 patients who have both tDNA (NGS or routine methods) and cfDNA NGS results. Using tDNA as reference, we found that the

TABLE 3 Mutation detection by NGS in tissue and plasma samples in discordant sample pairs

Sample ID	Tumor stage	cfDNA concentration (ng/mL)	Mutation detected by NGS in tissue (MAF)	Mutation detected by NGS in plasma (MAF)	Sequencing coverage in plasma	dPCR validation results in plasma samples (frequency)
21	IV	20	*EGFR* exon 19 del (46.3%)	NEG	1036.77	NEG
22	NA	12.28	*EGFR* L858R (28%) / *MET* gain	*EGFR* L858R (1.8%)	725.75	
23	NA	9.27	*EGFR* L858R (33%) / *MET* gain	*EGFR* L858R (1.3%)	844.77	
24	IIIa	8.2	*MET* gain	NEG	906.76	
25	IIIb	19.6	*KRAS* G12V (6.9%)	NEG	1034.87	NEG
26	IV	11.66	*EGFR* L858R (8.7%)	NEG	833.72	NEG
27	NA	12.52	*NRAS* Q61R (3.3%)	NEG	625.67	
28	NA	14.8	*HER2* gain	NEG	922.53	
29	IV	24.44	*EGFR* L858R / *KRAS* G13V (5%/4%)	*EGFR* L858R (18.0%)	1045.08	
30	IV	11.44	*EGFR* exon 19 del (67.4%)	NEG	1212.46	NEG
31	IV	16.99	*ALK-EML4* (0%)	NEG	1273.5	
32	IV	7.75	*EGFR* exon 19 del (19.2%)	NEG	1013.22	NEG
33	NA	27.17	*KRAS* G12C / *PIK3CA* H1047R / *EGFR* L861Q (23.7%/28.9%/18%)	NEG	1236.44	*KRAS* G12C NEG
34	NA	5.67	*EGFR* L858R / *KRAS* G12A (25%/3.2%)	*EGFR* L858R (0.1%)	784.27	
35	IV	20.84	*ALK-EML4*	NEG	1530.25	
36	NA	11.76	*EGFR* exon 19 del / *KRAS* G12R (14.6%/1.2%)	*KRAS* G12R (0.2%)	1276.93	
37	IV	15.46	*EGFR* L858R (2.1%)	NEG	1008.08	NEG
38	IV	37.69	*KRAS* G12C (7.3%)	NEG	1392.36	NEG
39	NA	11.2	*EGFR* exon 19 del (18.8%)	NEG	829.55	NEG
40	NA	14.7	*EGFR* exon 19 del / *EGFR* L 858R (21.7%/27.3%)	NEG	863.7	*EGFR* L858R POS, (0.0049)
41	NA	11.07	*EGFR* exon 19 del (15.6%)	NEG	1114.53	

Abbreviations: MAF, minor allele frequency; NA, not available; NEG, negative; POS, positive.

cfDNA NGS test exhibited high specificity for the detection of *EGFR* L858R (100%; 51/51), *EGFR* exon 19 del (100%; 55/55), *KRAS* G12X (100%; 57/57), and *ALK* rearrangement (98.3%; 59/60). The assay had a sensitivity of 76.2% for *EGFR* L858R (16/21), 58.8% for *EGFR* exon 19 del (10/17), 75.0% for *KRAS* mutations (6/8), and 50% for an *ALK* rearrangement (2/4) (Table 4).

3.7 | dPCR validation

All 21 patients with discordant results were positive on tumor and negative on plasma. To exclude FN results generated by NGS in plasma samples, of which most were SNVs and indels, dPCR was performed to 10 samples. We found one sample was positive using dPCR for *EGFR* L858R with a mutant allele frequency of 0.49%. Other discordant samples were not validated by dPCR because of sample insufficiency. Further, CNVs and rearrangements could not be validated because methods such as quantitative PCR require a significant amount of DNA, which is not available when using cfDNA.

4 | DISCUSSION

In this study, we evaluated NGS performance using a Beijing Genomics Institute (BGI) lung cancer panel on the Ion Proton Platform for the detection of gene mutations in cfDNA. To our knowledge, this is the first study to assess the performance of the Ion Proton for the detection of four distinct types of variants in cfDNA. Our assay had an overall sensitivity of 53.2% when all detected mutations were considered. Mutations in plasma samples were all identified in tumor samples, whereas 24 mutations identified in tumor tissue samples were FN using plasma. Among the 24 FN mutations, we found that 11 were SNVs, 7 were *EGFR* exon 19 deletions, 4 were CNVs, and 2 were rearrangements. Further, material was available for dPCR in 10 samples. Analysis of plasma samples with dPCR only confirmed the presence of a mutation in one case, which was *EGFR* L858R with a mutant allele frequency of 0.49%. Our findings using dPCR in NGS-negative plasma samples revealed that the variants found in tumor were likely either not shed into the circulation or were present at an ultra-low level that

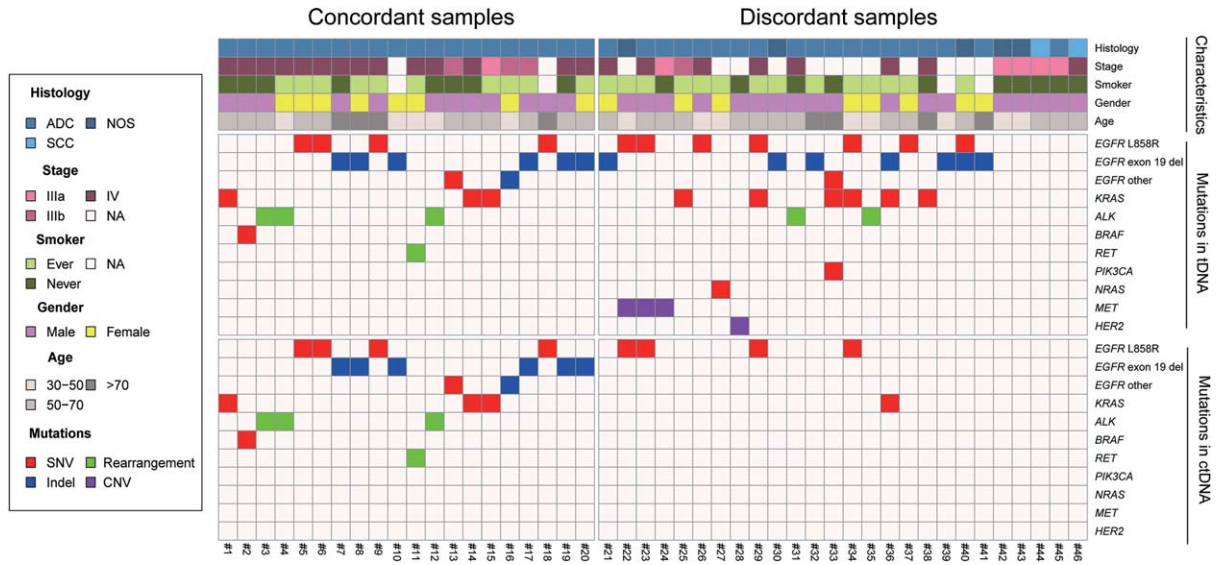

FIGURE 2 Mutation detection by tDNA and ctDNA NGS. Patients were categorized based on histology type, stage, smoking history, gender, and age (top). The mutation type (SNP, indel, CNV, rearrangement) identified in tDNA (middle) is compared with that found in plasma ctDNA (bottom). Samples without mutations are shown at the far right [Color figure can be viewed at wileyonlinelibrary.com]

was beyond our limit of detection, which was set at 1% for SNVs and indels. Among the 21 discordant samples, we found in addition to *EGFR* mutations that four samples had co-occurring mutations in tumor tissues, yet only *EGFR* L858R was detected in matched plasma samples. The two cases with co-occurrence of *EGFR* and *KRAS* mutations in tumor tissue are rare, and NGS results were confirmed in tumor tissue by dPCR. Furthermore, we also did not detect the four cases of *MET* amplification in cfDNA that were found by NGS using tDNA. *MET* amplification is associated with gefitinib resistance in patients with lung cancer and MET targeted therapy has received preclinical and clinical trials.[16,17] Identifying *MET* amplification in patients with lung cancer could aid the development of personalized medicine by identifying those patients who will benefit from treatment with MET inhibitors. For the failure of detecting CNV in our study, there are two potential explanations. On the one hand, the successful detection of alteration in cfDNA is highly depending on the proportion of tumor cell in the sample to be tested. As the actual frequency of CNV in cfDNA is unknown, a weak signal or fragmentation of the data could cause CNV signals to be masked by the intrinsic "background" fluctuation of the data. On the other hand, we used a targeted NGS panel that mainly designed to test SNVs, indels and rearrangements, while detection for CNVs was lack of optimization. Although algorithm for plasma CNV detection using WGS (whole genome sequencing) or WES (whole exome sequencing) have been reported in several places;[14,18–20] however, the robust detection of amplifications is limited by the context of low ctDNA fractions, and therefore tissue-based testing is considered to be a preferred reliable method for the detection of CNV.[9,21,22] Future efforts towards improving CNV detection using a targeted NGS panel in plasma are still needed.

Several studies, which have mostly focused on *EGFR* mutations, have reported different sensitivities for the detection of gene mutations in plasma cfDNA using different detection approaches. In one study, the sensitivity and specificity for *EGFR* mutation detection using droplet digital PCR (ddPCR) were reported to be 81.82% and 98.44% respectively,[23] whereas another study found a sensitivity of 65.7% and a specificity of 99.8% with Scorpion ARMS using the Therascreen kit (Qiagen),[24] which is the most frequently used method. The sensitivity of detecting *EGFR* mutations in plasma cfDNA using an NGS-based approach was comparable to that found using other methods. A meta-analysis of 27 studies using NGS reported a sensitivity and specificity of 62.0% and 95.9%, respectively.[10] In our study, we found a sensitivity and specificity of 68.4% and 100%. However, a limitation of ddPCR

TABLE 4 Sensitivity, specificity, and concordance rate of NGS ctDNA assay

Genetic alteration	*n*	Sensitivity (%)	TP	FN	Specificity (%)	TN	FP	Concordance rate (%)	Kappa (95% CI)
EGFR L858R	72	76.2	16	5	100	51	0	93.1	0.819 (0.669-0.969)
EGFR EXON 19 del	72	58.8	10	7	100	55	0	90.3	0.685 (0.475-0.895)
KRAS G12X	65	75.0	6	2	100	57	0	96.9	0.840 (0.625-1.000)
ALK rearrangement	64	50.0	2	2	98.3	59	1	96.9	0.547 (0.092-1.000)

Abbreviations: FN, false negative; TP, true positive.

TABLE 5 Genomic alterations associated with targeted therapies

Gene	Geneticalteration	Targeted therapy	Sensitive (S) or resistant (R)	Frequency (%)
All gene(s)				39
EGFR	L858R	Erlotinib, Gefitinib, Afatinib	S	16 (41.0)
	L861Q	Erlotinib, Gefitinib, Afatinib	S	1 (2.6)
	EXON 19 DEL	Erlotinib, Gefitinib, Afatinib	S	11 (28.2)
	EXON 20 INS	Erlotinib, Gefitinib, Afatinib	S	1 (2.6)
KRAS	G12X	Erlotinib, Gefitinib	R	6 (15.4)
ALK	*EML4-ALK*	Crizotinib, Ceritinib, Alecensa	S	2 (5.1)
RET	*RET-KIF5B*	Crizotinib	S	1 (2.6)
BRAF	V600E	Vemurafenib, Dabrafenib	S	1 (2.6)

and ARMs is the inherent inability to multiplex and test multiple biomarkers simultaneously. NGS has been intensively adapted for routine clinical use for personalized therapy.[25] A major advantage of this technology is the simultaneous testing of multiple clinically actionable gene mutations in a single assay that is both timely and cost-effective. Some more sensitive NGS-based detection method for liquid biopsy have been developed in recent years, but limitations exist for their widespread adoption in clinical application. For instance, Cancer Personalized Profiling by deep Sequencing (CAPP-SEQ), as a typical example for hybridization based methods, is expensive due to low percentage of reads mapping on-target. Oncomine, as an example for multiple PCR based technique, is unable to detect DNA rearrangements. Thus, a highly sensitive technique with affordable costs and ability to detect all major classes of mutations is needed for abroad application of liquid biopsy in routine clinical settings.

In the present study, 21 sample pairs were discordant with all alterations identified in specifically tDNA. The discordance may be due to the low concentration of ctDNA found in plasma, which results in a lower mutation frequency in plasma cfDNA than in tissue DNA. The results of dPCR partly supported this our assumption. Among 10 cases with mutations in tDNA only, merely one plasma was found positive using dPCR with a mutant allele frequency of 0.49%, indicating the mutations were either not shed into blood or the concentration of the mutation is below the detection limit of our method. In addition, our assay may not be sufficiently sensitive to detect very low frequency alterations that may be more readily identified by testing tissue, especially towards detection of CNVs and gene arrangements. In our matched sample pairs, four discordant samples have either a *MET* or *HER2* CNV in tDNA, but none was detected in the corresponding cfDNA samples (Supporting Information Figures S3 and S4).

Previous research has shown that the proportion of ctDNA in plasma varies by tumor size, disease stage, and the number of metastatic sites.[26,27] Similarly, our results show that M1 stage patients have higher concordance than M0 stage patients. These results suggest that the amount of cfDNA may reflect tumor burden and could be a biomarker for disease monitoring.

Clinically actionable alterations found in plasma cfDNA can also assist in guiding the selection of targeted therapies. Actionable alterations included in guidelines from the NCCN, AMP, ASCO, and CAP were found in 39 patients with lung cancer (54.2%) in our cohort (Table 5). Furthermore, *EGFR* alterations (L858R and exon 19 del) were the most frequent alterations found in cfDNA from 26 plasma samples in our study, both of which are actionable with approved EFGR TKIs (erlotinib, gefitinib, and afatinib).[3,6,28] Previous studies reported that *EGFR* mutations are predominantly observed in nonsquamous NSCLC, in patients who are nonsmokers, and in Asian patients.[29] In the current study, we observed a putatively significant difference between *EGFR* mutations and smoking status ($P = .057$). Additionally, drug resistance is a major clinical problem in patients responsive to EGFR-TKIs. An understanding of the molecular mechanisms responsible for acquired resistance is crucial including the development of primary resistance and the role of acquired mutations in *EGFR* (T790M and the exon 21 insertion) and in other oncogenes, such as *BRAF*, *PIK3CA*, *HER2*, and *MET*.[30,31] In our study, one *EGFR* exon 21 insertion and six *KRAS* activating mutations detected in cfDNA were previously associated with either a much lower response or acquired resistance to TKIs. No other acquired resistance mutation was found in our patient cohort. For inclusion in our study, these patients did not receive EGFR TKI-based therapy, which demonstrates the urgent need to provide a cfDNA test for actionable mutations early to patients with lung cancer to ensure they receive effective and timely personalized treatment.

In conclusion, we successfully developed a ctDNA NGS assay and analysis pipeline that can be used to detect cancer-related mutations and rearrangements in plasma samples obtained from patients with advanced NSCLC. Our results showed high concordance in the detection of clinically actionable alterations found in tumor tissues and paired plasma samples. In addition, we found that concordance rate is associated with cancer stage and tumor aggressiveness. Our ctDNA NGS test has great prospects for mutation profiling in patients who are unable to provide a histological sample, and may be a less invasive, more comprehensive, and cost-effective approach to identify targeted therapeutic strategies and monitor treatment response in patients with lung cancer.

ACKNOWLEDGMENTS

We appreciate the cooperation of all the patients and their families who contributed to this study. This project was supported by the National Key R&D Program of China (2016YFC0905400); the Science and Technology Planning Project of Guangdong Province, People's Republic of China (grant numbers 2007B031515017 and 2008A030201024); the Science and Technology Planning Project of Guangzhou, People's Republic of China (grant numbers 2007Z1-E0111 and 2007Z3-E0261); the Guangzhou Health and Medical Collaborative Innovative Major Special Projects (201400000001-2); special foundation for High-level Talents of Guangdong (2016TX03R171); Guangzhou Science and Technology Project (201400000004-5); Pearl River Nova Program of Guangzhou (No. 201506010065); and the Natural Science Foundation for Distinguished Young Scholars of Guangdong Province (No. 2014A030306013).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Han Liu http://orcid.org/0000-0002-2465-8814
Di Shao http://orcid.org/0000-0002-1485-2074

REFERENCES

[1] Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin*. 2014;64:9–29.

[2] Chen W, Zheng R, Zeng H, Zhang S. Epidemiology of lung cancer in China. *Thorac Cancer*. 2015;6:209–215.

[3] Maemondo M, Inoue A, Kobayashi K, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N Engl J Med*. 2010;362:2380–2388.

[4] Sequist LV, Yang JC-H, Yamamoto N, et al. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with *EGFR* mutations. *J Clin Oncol*. 2013;31:3327–3334.

[5] Solomon BJ, Mok T, Kim D-W, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med*. 2014;371:2167–2177.

[6] Zhou C, Wu Y-L, Chen G, et al. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol*. 2011;12:735–742.

[7] Ettinger DS, Wood DE, Aisner DL, et al. Non-small cell lung cancer, Version 5.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2017;15:504–535.

[8] Vanderlaan PA, Yamaguchi N, Folch E, et al. Success and failure rates of tumor genotyping techniques in routine pathological samples with non-small-cell lung cancer. *Lung Cancer*. 2014;84:39–44.

[9] Lanman RB, Mortimer SA, Zill OA, et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. *PLoS One*. 2015;10:1–27.

[10] Qiu M, Wang J, Xu Y. Circulating tumor DNA is effective for the detection of EGFR mutation in non-small cell lung cancer: a meta-analysis. *Cancer Epidemiol Prev Biomarkers*. 2015;24.

[11] Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. *Nat Med*. 2008;14:985–990.

[12] Shao D, Lin Y, Liu J, et al. A targeted next-generation sequencing method for identifying clinically relevant mutation profiles in lung adenocarcinoma. *Sci Rep*. 2016;6:22338.

[13] Liu J, Liu Z, Cheng S, et al. VHVT: an ultra-sensitive somatic mutation detection and performance assessment program. *J Clin Oncol*. 2017;35:1588–1588.

[14] Leary RJ, Sausen M, Kinde I, et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci Transl Med*. 2012;4:162ra154–162ra154.

[15] Xing B, Greenwood CMT, Bull SB. A hierarchical clustering method for estimating copy number variation. *Biostatistics*. 2007;8:632–653.

[16] Engelman JA, Zejnullahu K, Mitsudomi T, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science*. 2007;316:1039–1043.

[17] Pérez-Ramírez C, Cañadas-Garre M, Jiménez-Varo E, Faus-Dáder MJ, Calleja-Hernández MÁ. MET: a new promising biomarker in non-small-cell lung carcinoma. *Pharmacogenomics*. 2015;16:631–647.

[18] Ulz P, Thallinger GG, Auer M, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat Genet*. 2016;48:1273–1278.

[19] Chan KCA, Jiang P, Zheng YWL, et al. Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin Chem*. 2013;59:211–224.

[20] Xi R, Hadjipanayis AG, Luquette LJ, et al. Copy number variation detection in whole-genome sequencing data using the Bayesian information criterion. *Proc Natl Acad Sci USA*. 2011;108:E1128–E1136.

[21] Chung JH, Pavlick D, Hartmaier R, et al. Hybrid capture-based genomic profiling of circulating tumor DNA from patients with estrogen receptor-positive metastatic breast cancer. *Ann Oncol*. 2017;28:2866–2873.

[22] Ulz P, Belic J, Graf R, et al. Whole-genome plasma sequencing reveals focal amplifications as a driving force in metastatic prostate cancer. *Nat Commun*. 2016;7:12008.

[23] Zhu G, Ye X, Dong Z, et al. Highly sensitive droplet digital PCR method for detection of EGFR-activating mutations in plasma cell-free DNA from patients with advanced non-small cell lung cancer. *J Mol Diagn*. 2015;17:265–272.

[24] Douillard J-Y, Ostoros G, Cobo M, et al. Gefitinib treatment in EGFR mutated Caucasian NSCLC. *J Thorac Oncol*. 2014;9:1345–1353.

[25] Tran B, Brown AMK, Bedard PL, et al. Feasibility of real time next generation sequencing of cancer genes linked to drug response: results from a clinical trial. *Int J Cancer*. 2013;132:1547–1555.

[26] Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6:224ra24–224ra24.

[27] Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med*. 2014;20:548–554.

[28] Wu Y-L, Zhou C, Hu C-P, et al. Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-

small-cell lung cancer harbouring EGFR mutations (LUX-Lung 6): an open-label, randomised phase 3 trial. *Lancet Oncol*. 2014;15:213–222.

[29] Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst*. 2005;97:339–346.

[30] Camidge DR, Pao W, Sequist LV. Acquired resistance to TKIs in solid tumours: learning from lung cancer. *Nat Rev Clin Oncol*. 2014;11:473–481.

[31] Kobayashi S, Boggon TJ, Dayaram T, et al. *EGFR* mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med*. 2005;352:786–792.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Liu L, Liu H, Shao D, et al. Development and clinical validation of a circulating tumor DNA test for the identification of clinically actionable mutations in nonsmall cell lung cancer. *Genes Chromosomes Cancer*. 2018;57:211–220. https://doi.org/10.1002/gcc.22522